DOCUMENT RESUME

ED 272 920 CS 210 023

AUTHOR Yussen, Steven R.; And Others
TITLE Listening versus Reading in Monitoring Comprehension.
PUB DATE Apr 86
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Processes; Comparative Analysis; Higher Education; *Listening Comprehension; *Reading Comprehension; *Reading Processes; *Reading Research
IDENTIFIERS *Comprehension Monitoring; *Error Detection

ABSTRACT

Noting the differences in processing information by reading and by listening, two studies examined subjects' ability to detect errors in written and oral prose. In both experiments, college students were presented with four expository passages drawn from different written sources. All passages were approximately 300 words and 5 paragraphs long, with 4 paragraphs containing text disruptions specifically crafted for the study. Subjects read or listened to each passage, and noted any mistakes contained in each paragraph as it was being presented or during a 20-second pause between consecutive paragraphs. The general disruptions in the paragraphs consisted of statements of the main idea that contradicted one or more facts heard later, while specific disruptions consisted of single statements that were logically inconsistent or that contradicted common knowledge. Results of the two experiments did not appear to prove that spotting errors in a text is more difficult when listening than when reading. Mixed evidence suggested that in some cases listening is superior to reading. The two experiments did not permit drawing a clear inference about why listening superiority occurs sometimes, but two likely explanations are that (1) the pacing of the presentation is critical, and (2) subjects naturally process oral discourse more "deeply" than written discourse. (Tables of data are included, including two examples of disrupted paragraphs.) (HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED272920

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Listening Versus Reading in Monitoring Comprehension

Steven R. Yussen, Cecil Smith, Tina Huang, and Richard Wiermaa

Department of Educational Psychology

University of Wisconsin - Madison

Paper presented at the annual meeting of the

American Educational Research Association

San Francisco, California April, 1986

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Steven R. Yussen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS210023

During the past decade, a large number of investigations have been conducted to determine student's abilities to spot difficulties in prose (e.g. Markman, 1981; Baker and Brown, 1985; Garner, 1980; Garner and Anderson, 1982). The work has been referred to as error detection-emphasizing that the comprehender's task is to detect errors in the prose. Some of these studies have required students to read the prose (e.g. Baker and Anderson, 1981; Garner, 1980); other investigations have called on students to listen to the prose (e.g. Markman, 1979; Flavell, J.H., Speer, J.R., Green, F.L., and August, D.L. (1981)). As a number of scholars have noted, these two forms of processing information are different. So we must exercise caution in drawing inferences about how well comprehenders detect errors as a general matter in the face of the different ways the comprehension is being accomplished.

When listening, we generally cannot control the rate that the prose is presented and information unfolds in a linear, "beginning-to-end" type fashion. Moreover, there is often only a single opportunity to hear a specific piece of information, and without an opportunity to re-examine information "word-for-word", we must rely extensively on memory. By contrast, when we read we can control the rate that the prose is presented by speeding up or slowing down to accommodate the flow of our comprehension. And there are multiple opportunities to attempt processing of particular parts of the prose if we see fit simply by re-reading. Since we can move about the text backward and forward, the lock step "linearity" present in listening is avoided, and memory is less crucial since any particular "moment of reading" can be re-experienced if necessary.

The presentation offers impressions of the differences between comprehension monitoring in listening as opposed to reading contexts based on these considerations and the findings from two experiments recently completed

by the presenters with young adults. In the experiments, adults were asked to spot text difficulties in several expository lessons either while listening to them, reading them, or engaging in a combination of reading and listening. Across the two experiments, the listening experience was kept relatively constant and efforts were made to systematically manipulate the reading experience so that it would approximate the information constraints of listening to a greater and greater degree.

Overview of the Experiments

In both experiments, students enrolled in courses at the University of Wisconsin, were presented with four expository passages drawn from different written sources-an American Red Cross handbook on recreational activities, a book on home crafts, an encyclopedia, and a college-level study guide on philosophy.

The passages were selected because they covered a wide range of topics, were similar to the kind of reading students frequently encounter in school, and were challenging, based on the results of pilot work completed with students prior to the study. The passage topics were "Making Bread", "Modern Canoeing", "Folk Music" and "Fifth Century Greece". The original articles were edited to create uniformity of structure and length. All were approximately 300 words long with 5 paragraphs, each about 60 words. Four paragraphs contained text disruptions specifically crafted for the study. The third paragraph in each text had no discernible errors. The subjects' task was to read or listen to each passage, one at a time, and to note any mistakes contained in each paragraph as it was being presented or just shortly thereafter, when a 20 second pause between consecutive paragraphs permitted further time to write and explain any answer.

Given the extensive interest in the nature of text disruptions in the current literature, we should briefly mention the types of disruption worked into our paragraphs. They were of 2 kinds. General disruptions consisted of statements of the main idea or general point of the paragraph which contradicted one or more facts heard later. Specific disruptions consisted of single statements which were internally inconsistent or contradictory based on common knowledge and logic. Although the major purpose of the study was not to study different types of disruptions, per se, or to contrast different standards of comprehension as in Linda Baker's work, we expected subjects to find general disruptions more difficult to spot than specific ones. And we expected that subjects might be at a particular disadvantage with general disruptions as contrasted with specific ones, if they had to spot them while listening. This prediction was based on the reasoning that a) general text disruptions, to be spotted, require the contradictory elements to be held in short term memory, and b) listening effectively would seem to be especially dependent on short term memory.

An example of both types of disruption appears in Table 1.

Now let us turn to the details of each experiment.

Experiment 1

Forty eight students participated. They were volunteers from classes in Educational Psychology. They were randomly assigned to one of three conditions: 1) Listening only, 2) Reading only, or 3) Both listening and reading. Within conditions, subjects were exposed to all four passages and asked to identify any errors encountered, explain the error briefly, and rate the overall comprehensibility of each paragraph. Examples of a well constructed paragraph and an egregious one accompanied the instructions. For

each condition, two of the passages were presented once, the other two were presented twice. In the reading and reading and listening condition, booklets contained passages with paragraphs typed on a single page followed by a page for explanations and ratings of comprehensibility. Errors could simply be circled on the page of text. For listening, each paragraph was accompanied by a single page in the booklet for jotting down the error, explanations, and ratings of comprehensibility. The passages were heard on a standard portable cassette recorder (Craig Model J103) with 20 second delays between paragraphs. The audio cassette served to pace the reading in the reading and listening condition, but subjects paced themselves in the reading condition. We expected to find the following results:

1) Those passages presented for reading or listening twice should result in more errors detected than those passages presented for reading or listening once. Given two opportunities to spot an error, subjects should do better than if only given one opportunity.

2) Errors in which there was a specific disruption of the passage's meaning should be spotted more readily than errors in which there was a general disruption in meaning.

3) The condition in which subjects listened only should result in poorer detection of errors than the conditions in which subjects read or read and listened.

The principal measure employed to test for these effects was the correct identification of the error. Ratings of the comprehensibility of paragraphs (where flawed paragraphs should be perceived as less comprehensible) and explanations of any errors spotted, were considered to be supplementary measures of the effects predicted and described above.

The actual findings were as follows:

A three way analysis of variance was performed on the percentage of errors correctly detected by subjects. The independent variables included one that was between subjects--experimental condition (3) and two that were within subjects--the level of text disruption (2) and the number of times the subject was exposed to a passage (2). Table 2 displays the means associated with this analysis.

There was no statistically significant effect of experimental condition. The overall level of detecting errors was about the same regardless of whether subjects' listened only (2.38), read (2.30), or read and listened (1.94). It is noteworthy that not only was reading not better than listening, but listening actually was a little better (although not statistically so) than reading.

There was a statistically significant effect of being exposed to the passages for a different number of times, $F = 11.12$, d.f. = 1, 45, $p < .01$. Subjects benefitted from being exposed to passages twice (2.53 versus 1.81).

Finally, the effect of the type of disruption was significant $F = 14.15$, d.f. = 1, 45, $p < .01$. A greater number of general text disruptions were detected (2.41) than specific disruptions (1.94).

None of the interactions was statistically significant.

A similar analysis of variance was performed on the subjects' ratings of the comprehensibility of each paragraph in the passages. A 10 point rating scale was adopted with 10 defined as perfect comprehensibility and 1 defined as utterly incomprehensible. The means for this analysis are shown in Table 3. Keep in mind that here, lower means reflect lower perceived comprehensibility (which is consistent with perceiving the error as making the paragraph problematic to understand). The condition effect was not significant, although the effect of number of presentations ($F = 13.69$, d.f. =

1, 45, p <.001) and type of disruption (F = 32.10, $d.f.$ = 1, 45, p <.001) were significant. The main effects were due to similar patterns as in the preceding analysis. Following two presentations of the paragraphs, the comprehensibility of them was rated lower than following only one presentation. And the comprehensibility of paragraphs having general disruptions was rated lower than the comprehensibility of paragraphs having specific text disruptions.

As before, none of the interactions was statistically significant.

Brief discussion of the results:

We were quite surprised by these findings. There was no clear superiority of reading over listening. And to the extent that subjects varied in their error detection capabilities, they seemed to do better when engaged in a pure listening task, especially when they listened to the same passage presented twice. We were also surprised that subjects were better able to detect general text disruptions, when any effect surfaced at all.

Experiment 2

Since the findings were unexpected and since we had obtained no measures of other related cognitive performances for the subjects (thereby permitting us to rule out the significance of the group differences based on other known cognitive differences among the groups tested), we designed a second experiment aimed at replicating and extending the findings of the first one. In the second experiment, the passages were re-recorded in a professional sound studio to establish a better level of clarity on the tapes for a listening condition, and the written passages were read off an 80 column display of an Apple 2e, green, monochromatic monitor. The computerized display was to better control the subject's rate of exposure to the text.

This was done in one of two ways to be described momentarily in the presentation conditions. As before, there was a pure condition of 1) LISTENING only. It was identical to the condition in the first experiment with the exception that a new recording was played. Secondly, there was a 2) READING EXPERIMENTER CONTROLLED condition. Here, subjects saw one sentence of a paragraph scroll onto the screen from left to right at the same rate which it was read on the audio tape (for the LISTENING condition). The sentence then disappeared and the next one scrolled on automatically. So, for any given paragraph, the screen text duration was equated with the duration of the same paragraph on the audio tape (used in the LISTENING condition). Finally, there was a 3) READING SUBJECT CONTROLLED condition. Here, subjects saw one sentence of a paragraph scroll onto the screen, as in the second condition. However, advancing from one sentence to the next was controlled by the subject, who depressed any key on the computer terminal to generate the next sentence.

The purpose of shifting from ordinary reading of the passages on paper to reading the passages off the monitor of a computer terminal was to more precisely control the rate at which information was read and to establish more precise temporal equivalences between reading and listening. The subject controlled condition was added to circumvent the potential problem of some normal disruption of text comprehension created by attention to the visual novelty of "scrolling" text on a screen. By allowing subjects to look at each individual sentence a little longer than its aural (listened to) counterpart in this third experimental condition, subjects were being given extra time, which might work to their advantage in detecting errors.

Each subject was also administered three additional tasks to a) ascertain if the subjects randomly placed into the three groups were alike in some

cognitive respects, and b) to determine if individual differences in error detection performance are predictable from other measures of verbal ability. One task measured ordinary READING TIME, by giving subjects an 1150 word passage, drawn from a popular textbook on the art of writing, and asking them to read it as they would read most things each day. A second task measured general verbal ability-the VOCABULARY subtest of the WAIS, in which subjects are asked to define 35 words. The third task measured WORKING MEMORY for text, utilizing a measure modelled after one described by Daneman and Carpenter (1980, 1983). Each of these measures was administered approximately one week after the error detection task.

Means and standard deviations for the cognitive measures appear in Table 4. As can be seen, the three experimental groups are highly similar in their average reading times, vocabulary performance, and verbal working memory. A one way analysis of variance for each dependent measure revealed no significant differences among groups for any measure.

Means associated with the error detection task appear in Table 5.

A two way analysis of variance for presentation condition (3) and type of text disruption (2) yielded significant effects for both main effects and there was a marginally significant interaction. Table 5 shows the means associated with these effects. As can be seen, the mean level of detecting errors was higher in the listening condition than either of the reading conditions. Post hoc Tukey comparisons, showed that performances in the two reading conditions were not different from one another, but were each different from performance in the listening condition. And, as in experiment 1, subjects detected more of the general disruptions than specific disruptions. The marginally significant interaction appeared to be due to the relative absence of differences in level of detecting errors across

experimental conditions when one considers the specific text disruptions only (row 2 in Table 5), but large differences when the general disruptions are considered (row 1 in Table 5).

Correlational analyses among the three cognitive measures and the error detection performances within each of the experimental groups revealed very little. Of the 9 key correlations (each cognitive measure correlated with error detection performance separately for each of the 3 experimental conditions), only one yielded a statistically significant effect. Vocabulary performance yielded a Pearson $r = .45$ ($p<.05$) with error detection performance in condition 2 (READING EXPERIMENTER CONTROLLED).

General Discussion

From the results of the two experiments taken together, it does not appear that spotting errors in a text proves to be more difficult when listening to a text than when reading or reading and listening. There is mixed evidence to suggest that just the opposite might be true in some cases-that listening is superior to reading. The two experiments do not permit a clear inference to be drawn about why this listening superiority is sometimes manifested, but two likely explanations are that a) the pacing of the presentation is critical and that b) subjects naturally process "heard discourse" more "deeply" than written discourse. Perhaps, the specific rate created in some conditions (about 100-120 words per minute) proves to be a much more optimal rate for listening than for reading. Or perhaps college subjects, used to experiencing much important information to be mastered in oral (classroom) contexts, naturally adopt a deeper level of processing when they listen than when they read. The clear message for future research, then, is to explore different rates of presenting information to determine what is

most optimal and to consider varying instructions to subjects which may influence their depth of processing the material.

A brief comment is also in order for the consistent finding that general level text disruptions proved easier to spot than ones at a specific level. Our interpretation of this finding is that the particular disruptions one chooses can vary considerably for subjects in how subtle or obvious they prove to be. And our general level disruptions were apparently more obvious than the specific level ones. Unless striking demands are placed on short term memory-that is, making it difficult to hold the discrepant information in mind, then, the temporal unfolding of the errors is less critical in influencing performance than the nature of the errors, themselves. And, apparently, such difficult demands were not present in our passages.

REFERENCES

Baker, L. & Brown, A.L. (1984). Metacognitive skills and reading. In D.P. Pearson (Ed.), Handbook of Reading Research. New York: Longman.

Daneman, M. and Carpenter, P.A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19, 450-466.

Daneman, M. and Carpenter, P.A. (1983). Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 561-584.

Flavell, J.H., Speer, J.R., Green, F.L., and August, D.L. (1981). The development of comprehension monitoring and knowledge about communication. Monographs of the Society for Research in Child Development, 46, (5, Serial No. 192).

Garner, R. (1980). Monitoring of understanding: An investigation of good and poor readers' awareness of induced miscomprehension of text. Journal of Reading Behavior, 12, 55-64.

Garner, R. & Anderson, J. (1982). Monitoring-of-understanding research: Inquiry directions, methodological dilemnas. Journal of Experimental Education, 50, 70-76.

Markman, E.M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. Child Development, 50, 643-655.

Markman, E.M. (1981). Comprehension monitoring. In W.P. Dickson (Ed.), Children's Oral Communication Skills. New York: Academic Press.

Table 1

Examples of Disruptions

(General disruption-initial paragraph of "Greece" Passage)

The fifth century B.C. is considered one of the periods of lowest cultural development in the history of western man. Drama absorbed lyric poetry and created a form of literature unique in the ancient world. History reached its maturity, and philosophy revolutionized thought processes. Sculpture and architecture were so highly refined that they became the classical standards for future generations.

(Note: The accomplishments cited conflict with the opening statement that the century marks the lowest period of cultural development).

(Specific level text disruption-final paragraph of "Folk music" Passage)

A number of new types of folk music arose among North American settlers. One type was the immigrant song in which the settler sang of his homesickness and the hardships of a new life. Another musical tradition was the travelling songs of the Canadian fur traders who travelled by canoe through the Great Lakes region. Slave workers in the cotton fields of the old South carried on the ancient call-and-response patterns of African music, as did members of railway construction gangs who rarely sang such songs.

(Note: In the final sentence, there is a contradiction. Railway construction workers are said both to have carried on the singing of African music, and to have rarely sung such songs).

Table 2

Means and Standard Deviations for the Number of Errors Detected By Condition, Type of Error, and Number of Times the Passages Were Presented

Experiment 1

Type of Text Disruption and Number of Presentations	Reading n=16		Listening n=16		Reading and Listening n=16	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
General						
1 time	1.69[a]	(1.35)	2.25	(1.39)	1.94	(1.23)
2 times	2.38	(1.20)	3.31	(.87)	2.88	(.80)
Specific						
1 time	1.63	(1.36)	1.38	(1.14)	2.00	(1.10)
2 times	1.69	(1.13)	2.56	(.96)	2.38	(1.36)

[a]The range of scores was 0-4, based on each subject's performance in detecting errors (0-4) averaged across the 4 stories presented.

Table 3

Means and Standard Deviations for the Paragraph Ratings According to Condition, Type of Error Present in the Paragraph, and Number of Times the Paragraphs Were Presented

Experiment 1

Type of Text Disruption and Number of Presentations	Reading n=16		Listening n=16		Reading and Listening n=16	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
General						
1 time	7.10[a]	(2.37)	7.15	(1.54)	7.28	(1.94)
2 times	5.88	(1.55)	6.45	(1.43)	6.30	(1.10)
Specific						
1 time	7.05	(1.89)	8.11	(1.31)	8.09	(1.08)
2 times	7.22	(1.23)	7.39	(1 31)	7.20	(1.32)

[a]The rating scale was from 1 (incomprehensible) to 10 (perfectly comprehensible).

Table 4

Means and Standard Deviations for Three Cognitive Measures for Each of the Experimental Groups

Cognitive Measures	Experimental Condition: Listening		Reading - Exp. Controlled		Reading - S. Controlled	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Reading Time (secs)	226	(22)	223	(25)	218	(28)
Vocabulary (WAIS Score)	39.05[a]	(6.46)	37.05	(6.86)	39.50	(6.27)
Verbal Working Memory	2.70[b]	(.92)	2.85	(.99)	3.00	(.86)

[a]The range of scores on the WAIS is from 0 to 70.

[b]Subjects received a score for the longest series of words they could recall. Hence 2=2 words, 3=3 words, etc. This is a demanding task, since the words recalled are the last ones in a series of complex sentences that must be read prior to recall.

Table 5

Means and Standard Deviation for the Number of Errors Detected by Condition and Type of Error

Experiment 2

Type of Text Disruption	Experimental Condition: Listening $\bar{X}$	Listening S.D.	Reading - Exp. Controlled $\bar{X}$	Reading - Exp. Controlled S.D.	Reading - S. Controlled $\bar{X}$	Reading - S. Controlled S.D.
General	5.85[a]	(1.46)	4.00	(1.65)	4.55	(1.82)
Specific	3.10	(1.62)	3.00	(1.56)	3.05	(1.93)

[a]There were 8 possible errors to detect (4 passages x 2 errors of each type per passage-general or specific).